DOCUMENT RESUME

ED 131 417 CS 003 015

AUTHOR Bayer, Maria L.
TITLE Primary Grade Understanding Vocabulary as Measured by Orally Administered Basic Word Vocabulary Test.
PUB DATE Oct 76
NOTE 65p.; M.Ed. Thesis, Rutgers, The State University of New Jersey

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Basic Vocabulary; Elementary Education; Masters Theses; Reading Research; Reading Tests; Test Construction; Test Validity; *Vocabulary Development
IDENTIFIERS *Basic Word Vocabulary Test; BWVT

ABSTRACT

The Basic Word Vocabulary Test (BWVT) was extended downward by administering it orally to 96 pupils selected randomly from grades one through four. Estimates of aural understanding vocabulary at each grade were obtained and were found to be much lower than results from previous studies, due to the strict criteria of the BWVT. Correlations with the Gates-MacGinitie reading tests were .74 for vocabulary and .54 for comprehension. The correlation with Kuhlman-Anderson IQ scores was .29. Results were interpreted as supporting the feasibility of using the BWVT as an orally administered measure. (Author/AA)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

PRIMARY GRADE UNDERSTANDING VOCABULARY AS MEASURED BY ORALLY ADMINISTERED BASIC WORD VOCABULARY TEST

AN ABSTRACT OF A THESIS

SUBMITTED TO THE FACULTY

OF THE GRADUATE SCHOOL OF EDUCATION

OF

RUTGERS

THE STATE UNIVERSITY OF NEW JERSEY

BY

MARIA L. BAYER

IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE

OF

MASTER OF EDUCATION

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY

Maria L. Bayer

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

COMMITTEE CHAIRPERSON: Edward B. Fry, Ph.D.

NEW BRUNSWICK, NEW JERSEY

OCTOBER, 1976

CS 003 015

Abstract

This study was an attempt to extend the Basic Word Vocabulary Test (BWVT), developed by Dupuy (1974), downward by administering it orally. Four questions were investigated:

1. How large are the aural understanding basic word vocabularies of children in grades one through four?

2. Do significant differences exist among grade levels?

3. Do significant differences exist between the orally administered BWVT scores and the reading BWVT scores of third and fourth grade children?

4. What correlations exist between the oral scores and IQ, and between oral scores and the Gates-MacGinitie Reading Test results?

For investigation of these questions, 96 students (24 per grade level) were randomly selected from grades one through four in a Central New Jersey township school. The BWVT was administered orally to all four grades and as a reading test to third and fourth grades.

The estimates of aural understanding vocabulary size were first grade: 1,780 basic words; second grade: 2,330 basic words; third grade: 3,780 basic words; and fourth grade: 4,000 basic words. These results are much lower than results of previous studies.

A one way analysis of variance and Scheffe's tests on contrasts were performed on the oral means. There was a main effect of grade ($F = 32.06$, $df = 3/92$, $p < .01$), and all pairwise comparisons were significant at the .01 level.

Significant differences between the oral and the reading means for third and for fourth grade were verified by t Tests. The t values of 6.9 for third grade, and 6.3 for fourth grade were both significant at the .001 level.

The correlation between the oral BWVT and 1) Gates-MacGinitie vocabulary was .74, ($p < .01$), 2) Gates-MacGinitie comprehension was .59, ($p < .01$), and 3) Kuhlman-Anderson IQ was .29, ($p < .05$). The low .29 correlation may have been due to the IQ data being old.

The experiment showed the feasibility of administering the BWVT orally. The low vocabulary size estimates are more in accordance with the early vocabulary size estimates than with estimates dependent on the Seashore-Eckerson test. This is due to Dupuy's strict criteria in defining a basic word.

PRIMARY GRADE UNDERSTANDING VOCABULARY

AS MEASURED BY ORALLY ADMINISTERED

BASIC WORD VOCABULARY TEST

A THESIS

SUBMITTED TO THE FACULTY

OF THE GRADUATE SCHOOL OF EDUCATION

OF

RUTGERS

THE STATE UNIVERSITY OF NEW JERSEY

BY

MARIA L. BAYER

IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE

OF

MASTER OF EDUCATION

NEW BRUNSWICK, NEW JERSEY OCTOBER 1976

APPROVED: _____
EDWARD FRY

_____
MARJORIE ARNOLD

_____
JOHN SANTA

DEAN: _____
MILTON SCHWEBEL

ACKNOWLEDGMENTS

I would like to express my appreciation to Dr. Edward Fry, Director, Reading Center, for his encouragement and helpful suggestions during the preparation of this study.

I would also like to thank Dr. Marjorie Arnold for her comments and questions which made the writing easier, and Dr. John Santa for his advice on some of the statistical writing.

Gratitude is expressed to Mr. Rubin Feldstein, principal, Irene E. Feldkirchner school, and to all the teachers for their cooperation in the testing procedure.

I would like to especially thank my husband Douglas, and son David, for their help and support during the time I paid more attention to "the thesis" than to them.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	ii
LIST OF TABLES	v
Chapter	
I. STATEMENT OF THE PROBLEM	1
Background of the Problem	1
Statement of the Problem	5
Importance of the Problem	6
Definition of terms	7
Limitations of the Study	8
Overview of the Study	9
II. REVIEW OF THE LITERATURE	10
Trends	10
Methods	13
Problems	15
Definition of "word"	15
Definition of "know"	17
Definition of vocabulary	19
Errors in sampling	20
Tests	23
Summary	24

TABLE OF CONTENTS (continued)

Chapter		Page
III.	PROCEDURE	25
	Population	25
	Instrument	27
	Pilot Study	31
	Design	31
	Data Collection	32
	Data Analysis	33
IV.	FINDINGS AND DISCUSSION	35
	Presentation of Data	35
	Significant Differences	37
	Correlations	38
	Discussion	41
V.	SUMMARY AND CONCLUSIONS	47
	Summary	47
	Conclusions	49
	Suggestions for Further Research	50
BIBLIOGRAPHY		52
APPENDIX		
A.	Sample Basic Word Vocabulary Test	55

LIST OF TABLES

Table		Page
1.	Some Estimates of English Vocabulary Size, by Method Used.	12
2.	Means for Oral BWVT, Reading BWVT, Vocabulary Size, Ages, IQ, and Gates-MacGinitie Reading Test, by Grade Level . . .	36
3.	Analysis of Variance Among Grade Levels . . .	39
4.	Analysis of Mean Difference Among Grade Levels using Scheffe's Tests on Contrasts . .	39
5.	*t* Test Between Oral Administration and Reading Administration of the BWVT for Third and Fourth Grades	40
6.	Correlation Matrix of Oral BWVT Raw Scores, Gates Vocabulary Grade Scores, Gates Comprehension Grade Scores, and Kuhlman-Anderson IQ Scores.	41

CHAPTER 1

THE PROBLEM

An important phase of reading instruction is vocabulary development. A vast amount of literature exists on how to develop vocabulary and many investigators have attempted to measure the vocabulary size of children and adults. However, when vocabulary size has been studied and results from similar groups compared, the results are not consistent. Some vocabulary size estimates are 10 times the size of estimates made earlier (Lorge & Chall, 1963).

Lorge and Chall (1963) concluded that the reported vocabulary size variations were due to the differences in methodologies used by various investigators. Two major methods for estimating vocabulary size are the sampling method and the use method. With the sampling method, words can be systematically sampled from a dictionary or Thorndike's word counts. Usually every nth word on predetermined pages, or every nth word in a list is selected for testing. Another technique in sampling depends on using a table of random numbers to locate the words in a dictionary for testing. The vocabulary size estimate is calculated by multiplying the sample number correct by the ratio of the source size to the sample size. (Suppose a

sample of 100 words was drawn from a source containing 10,000 words. A subject might get 80 of the 100 sample words correct. The subject's vocabulary size would be estimated at 8,000 words, since 80 X 10,000/100 = 8,000.) This method limits the total number of words that a person knows to the size of the source, and depends on the person recognizing (usually by reading) the words. Seashore and Eckerson (1940) used the dictionary sampling method in determining the vocabulary size of college students.

The use method depends on a person's producing (rather than recognizing) words either through speech or writing. The vocabulary size is estimated by totaling the different words used. A limitation of this method is that people will not use all the words they know in a given situation, as pointed out by Seashore and Eckerson (1940). One investigator who used this method was Rinsland (1945). He estimated children's vocabulary size based on their writing.

With both methods there are difficulties, such as the definition of a "word". If "bat" is counted as a word, how is "bats" to be counted? One may argue that it is simply a plural and should not be counted as an additional word, or investigators may count it as a separate word. Another difficulty with counting a "word" arises when the meanings are considered. Is "bat" to be counted as one word or two words when a person knows it is an animal as well as a piece of baseball equipment? In just the unit of measurement, there is a source of vocabulary size variation, due

to lack of agreement on the definition of "word".

Another difficulty is the definition of "know". There is a whole range of "know" from being able to define the word easily, to being uncertain, to not even knowing if the collection of letters is a word (Dale, 1965). One may be able to define "batted" the ball, but is it as easy to define "batted" an eye? Does "enervate" mean to increase or to decrease strength? What does "praedial" mean? One investigator may require choosing one of three alternatives to define a word. Another may require the subject to produce a synonym. Not only is the criteria for "know" different, introducing another source of size variation, but there are two different vocabularies being tested.

Thus, another problem area is the designation of vocabulary type. It may be general knowledge that people have different-sized vocabularies for speaking, writing, reading, and aural understanding. But the designation of which specific vocabulary was measured is often omitted in a general statement like, "First-graders know 12,000 words." There is quite a difference if this number refers to a speaking, a reading, or a writing vocabulary. Children, particularly in primary grades, will have smaller writing vocabularies than understanding vocabularies because they are just learning to read and write. As people mature, the vocabulary sizes change with respect to each other, .i.e the reading vocabulary becomes larger than the speaking vocabulary. One may understand "enervate" within its printed

context, but not use it in speaking. Obviously, measuring vocabulary size is a task requiring several considerations. It was Lorge and Chall's belief that different methodologies with the aforementioned difficulties caused the variance in vocabulary size estimates.

Similarly, Dupuy (1974,p.2) felt the inconsistent vocabulary size results were due to the investigators' failure to define clearly four aspects of measurement:

(1) definition of the unit of measurement--the word

(2) estimates of the word population

(3) basis for sampling, e.g., the size of the dictionary or the nature of the use situation from which the sampling for the test was taken, and

(4) criteria used in determining word knowledge.

With these points in mind, Dupuy constructed the Basic Word Vocabulary Test (BWVT). It measures both absolute vocabulary size (total number of words that an individual knows) and relative vocabulary size (size of the individual's vocabulary in relation to vocabulary size of other groups).

Dupuy developed criteria for each aspect of measurement. First, the estimate of the word population was taken from Webster's Third New International Dictionary. Main entries from one of three columns on every tenth page were counted. Only one of all similar homographs was counted; abbreviations were counted, but not affixes. The process yielded a rounded population estimate of 240,000 words,

from which a 1-percent sample was taken. This sample was further reduced by Basic Words having to meet these requirements: they had to be main entries in all four specified major dictionaries, they had to be single word forms--not compounds, hyphenated forms, or abbreviations-- and they could not be designated as a technical, foreign, slang, or archaic word.

A word was "known" when it was correctly chosen from five alternatives on a multiple choice reading test. Strict criteria were also used in the construction of the correct responses and the distractors for each item. Items were ranked according to difficulty determined in pretesting situations.

The BWVT is a multiple choice test whose range is third grade through Ph.D. level. A problem with the test is that it is too difficult for first and second graders, since they lack the necessary decoding skills to read the test items. How can their vocabulary be assessed then? In discussing testing the word knowledge of those with limited reading ability, Lorge and Chall (1963, p. 149) state that "...the problem may be avoided by administering the test orally."

Statement of the Problem

The present study is designed as a downward extension of the Basic Word Vocabulary Test by administering it orally. The major questions to be studied are

(1) How large are the aural understanding basic word vocabularies of children in grades one through four?

(2) Do significant differences exist among grade levels?

(3) Do significant differences exist between the oral BWVT scores and the reading BWVT scores of third and fourth grade children?

(4) What correlations exist between the oral scores and IQ, and between oral scores and the Gates-MacGinitie Reading Test results?

Importance

The downward extension (oral administration) of the BWVT provides a needed tool that measures both 1) the absolute (cardinal number result) and 2) the relative (ordinal value result) aural understanding vocabulary sizes of primary grade children. At present, there is no other current instrument which can give both types of measurement.

It is also important to understand that the unit of measurement--the basic word--is a very carefully defined unit. There should be no confusion as to whether compounds, technical words, or variant forms are included, thus inflating the vocabulary size.

The main purpose in using the BWVT is to gather accurate baseline data on individuals. By assessing the absolute and the relative size of a student's vocabulary, the

examiner would gain knowledge about the individual, and how he compares to the group. The individual's growth patterns could be monitored.

As a diagnostic tool, the BWVT should be useful for indicating which individuals need extra help in vocabulary building. Teachers could then group these children for instruction.

Group scores could provide a basis for recommendation in curriculum change, e.g. should more stress be placed on vocabulary instruction? Which words should be included in primary reading lessons?

Definition of Terms

Absolute vocabulary size. The actual size of one's vocabulary expressed in a cardinal number, e. g., 12,700 words.

Basic word. A word that meets Dupuy's five criteria of being 1) a main entry in all four of these dictionaries: Webster's Third New International, Random House, World Book, and Funk and Wagnell's; 2) a single form, not compound or hyphenated; 3) not a proper name or abbreviation; 4) not foreign, archaic, technical, or slang; and 5) not a derived or variant form.

Relative vocabulary size. The size of one's vocabulary relative to another individual or group, expressed in a rank value, e.g. the 50th percentile.

Vocabulary. The collection of words for which people

have meaning, especially the words understood when people hear them.

Word. When "word" is used alone, it refers to a lexical entry in the dictionary.

Limitations

One of the problems in assessing vocabulary size is determining when a word is known and thus counted in one's vocabulary. In the BWVT, examinees have to be able to demonstrate only one meaning of the basic word by selecting one of five choices. This procedure does not allow an examinee to show how many meanings of the basic word he knows. Thus, the BWVT measures vocabulary breadth, not depth. It must also be understood that the BWVT, used in this way, measures aural understanding vocabulary size, not reading, speaking, or writing vocabulary sizes.

Another limitation of the present study is the number of subjects over the one to four grade range. Due to the time involved in testing small groups of four subjects in first and second grades, the population had to be restricted. Only 24 subjects per grade level , for a total of 96 subjects, were used.

Data for IQ are above the mean and data for reading levels are above grade levels. This indicates the population is above average, and the results should not be projected onto other groups.

Overview of the Study

Chapter Two is a review of the literature on vocabulary studies. The trends in research, different methods for estimating vocabulary size, and inherent problems are included.

Chapter Three is a description of the procedure for extending the BWVT downward and for obtaining the aural basic word vocabulary size of primary grade children. Descriptions of the population, Dupuy's test, pilot testing, data collection, and data analysis are included.

Chapter Four is a presentation of the data and discussion of the findings in relation to the literature. Data include the means of the oral scores by grade level and an analysis of variance of the means, t tests between oral and reading scores of the BWVT by grade level, and a correlation matrix of the oral scores, Gates-MacGinitie Reading Test scores, and the Kuhlman-Anderson IQ test scores.

Chapter Five is a summary and a presentation of the conclusions of this study. Suggestions for further research are included.

CHAPTER II

REVIEW OF THE LITERATURE

This experiment was an attempt to extend the Basic Word Vocabulary Test (BWVT) downward through oral administration. Given as a reading test, the instrument measures reading basic word vocabulary size, but given orally, it assesses the aural understanding basic word vocabulary size. The review of the literature includes the trends, methods, and problems of vocabulary size research, and comments on vocabulary tests. The trends section states when most of the studies were done and what is recent in the area. The methods section describes dictionary and frequency sampling as well as methods based on use. The problems section deals with the difficulty in defining a word, in defining "know" and testing it, with the need to define the kind of vocabulary being tested, and lastly, with errors in sampling.

Trends

Most vocabulary size studies were done in the nineteen-twenties to the nineteen-fifties: Terman(1916), M.E. Smith (1926), Dolch (1936), Seashore & Eckerson (1940), M.K.Smith (1941), Rinsland (1945), and Shibles

(1959). The early studies by Terman, M.E. Smith, and Dolch revealed primary grade children knew less than 4,000 words. The later studies by M.K. Smith and Shibles showed primary grade children knew 24,000 words or more. Controversy over the estimated size variations has continued for years. (Table 1 is a summary of the results of some previous studies in vocabulary size research.) Finally, Lorge and Chall (1963) analyzed the methods used and contended that the reasons for vocabulary size discrepancies lie in the procedures used to estimate size. These procedures are elaborated in the following Methods section. Table 1 is a summary of the results of some of the studies made in vocabulary size.

Little research has been done in the area of vocabulary size recently. Ames (1964) constructed a test, using the dictionary-sampling method, to find the understanding vocabulary size of first-grade students. His data include a mean vocabulary size of 12,456 words, with a standard deviation of 2,952 words. ERIC, Current Journals in Education, Education Digest, Psychological Abstracts, and Dissertation Abstracts were searched for studies made in the last five years in the areas of basic vocabulary, vocabulary studies, and vocabulary development. Dupuy's Basic Word Vocabulary Test is the only vocabulary size and development study undertaken within the last five years.

TABLE 1

SOME ESTIMATES OF ENGLISH VOCABULARY SIZE BY METHOD USED

AUTHOR(DATE)	VOCABULARY SIZE	AGE OR EDUCATIONAL LEVEL
Dictionary Sampling		
Ames (1964)	12,456	first graders
M.K.Smith (1941)	24,000	first graders
Shibles (1959)	26,000	first graders
Dupuy (1974)	1,532	third graders
Terman (1916)	3,600	third graders
M. K. Smith (1941)	44,000	third graders
Dupuy (1974)	4,419	seventh graders
Terman (1916)	7,200	seventh graders
Holley (1919)	8,478	seventh graders
Kirkpatrick (1907)	10,666	seventh graders
Seashore (1933)	65,000	college sophomores
Gerlach (1917)	85,300	college sophomores
Frequency Sampling		
M.E. Smith (1926)	2,562	first graders
Dolch (1936)	2,703	first graders
Use: Spoken		
Horn (1928)	3,145	one five year old
Use: Written		
Rinsland (1945)	5,000	first graders

Methods

Different methods of estimating vocabulary size include methods based on use, frequency-sampling, and dictionary- sampling (Lorge & Chall, 1963). Methods based on use means estimates are made on the number of words used in writing or speech. For example, Rinsland (1945) counted the words children wrote in compositions, stories, and letters. He found that first graders knew over 5,000 words. For spoken vocabulary measurement, M. Horn, in 1928, reported using a dictaphone and parent checklist to find that a five year-old girl knew over 3,145 words (Dale, 1956, p.115). Buckingham and Dolch (1936) used the free response method to determine what words children from grade to grade knew. They directed children to write down all the words that they could think of in 15 minutes. While the object of the study was to create a word list, Buckingham and Dolch also considered vocabualry size. At the end of first grade, children knew 2,800 words. With these methods, vocabulary size is limited to the words children use in the particular situations. Seashore (1940, p.19) states:

> In certain studies of children it is possible to record all of the words used in a variety of situations over a period of days until no new words appear except those learned during the period. This gives a measure of the number of words which the individual did use, but it does not take into account those which he could have used if necessary, nor those which he could only recognize.

Vocabulary size can also be estimated by frequency

sampling (Lorge & Chall, 1963). In this method, Thorndike's books of frequency word counts are used. Words are taken from each specified frequency level, and the subject is tested on them. The vocabulary is estimated by adding the number known at each level. For example, Thorndike's A Teacher's Word Book of 20,000 Words might be used to test a student. Words could be selected from the commonest thousand, second thousand, and third thousand frequency levels for testing. The student may get correct half the words from the commonest thousand, a quarter of the words from the second thousand, and none from the last. The subject's vocabulary size would be estimated at 750 words: 500 (1/2 X 1,000) + 250 (1/4 X 1,000) + 0 (0 X 1,000) = 750. In the frequency-sampling method, vocabulary size is limited by the number of words in the lists--10,000, 20,000, or 30,000. M.E. Smith (1926) found that first graders knew 2,562 words using the 10,000 word list.

The dictionary-sampling method involves selecting a dictionary, taking a sample of words from it, and testing a subject on the sample. To get the subject's vocabulary size, the number correct on the sample is multiplied by the ratio of the dictionary size to its sample size. For example, a sample of 100 words may be taken from a dictionary having 50,000 words. A subject gets 90 words correct on the sample. It is estimated that he would know 45,000 words since 90 x 50,000/100 = 45,000. This method limits the number of words a person can know to the size of the

dictionary from which the sample was drawn. Ames (1964) used this method in finding the mean first grade vocabulary size to be 12,456 words.

Ideally, an investigator could count all the words in the English language and test an individual to determine how many words he knows. The dictionary sampling method comes closest to this, but there are inherent problems.

Problems

Many problems exist in vocabulary size research. These include 1)defining a word, 2)defining "know" and testing it, 3)failure to designate which vocabulary is being tested, and 4) sampling errors. Each problem is discussed in turn.

Definition of "word". Since investigation of vocabulary size began, there has been no unanimous agreement on what exactly constitutes the basic unit of measurement--the word. Obviously, different units of measurement would yield variations in estimated vocabulary size. This is a problem in all methods used. Generally, most investigators accept a lexical entry in a dictionary as one word, and include with that one word the variant forms such as plurals, comparatives and superlatives, and participles. When the variant forms are counted separately, the vocabulary estimate would be larger.

This disparity in method is illustrated by three studies. First, Ames (1964) randomly selected entries from a

dictionary which were printed in boldface type. He did not include derived words or words formed from the basic words, which were in medium type and indented. Secondly, Seashore and Eckerson (1940) considered both basic and derived words in a dictionary. Basic words were typed in boldface print next to the margin, derived words were typed in medium print and indented. The lists were combined to make a third "total words" list. Thirdly, Seashore and Eckerson (1940, p.16)) cited Thorndike's rules for classifying words as variations of a word. A variation, not to be counted separately, could be one of these: plurals "s" and "ies", adverbs ending in "ly", comparatives and superlatives, verb forms ending in "s", "d", "ed", "ing", past participles ending in "n", and adjectives adding "n" to proper nouns.

Just these three examples of defining a word yield a different number of total words to be sampled. Seashore and Eckerson (1940, p. 16) stated their method of counting was different than Thorndike's:

> By counting all of the words in the dictionary which would be grouped with our basic sample according to Thorndike's definition we find that, in proportion, our estimate of three hundred seventy thousand two hundred sixty as the total number of words in the dictionary would be equivalent to a much smaller number of words according to Thorndike's definition.

Seashore and Eckerson's list would also be smaller (limited to their basic word list) if counted by Ames' method.

Dale (1965, p.898) followed Thorndike's rules in discussing counting inflectional forms and the age of the sub-

ject being tested. He counted a word as new "when the inflection is likely to make it markedly more difficult. Thus you may know like but not likely. This problem is greater for the younger child who has not generalized the regular endings." But just how one defines "markedly more difficult" is not described, and seems imprecise.

Definition of "know". What does it mean to "know" a word? Not only the depth of meaning but the degree to which one realizes that a group of sounds or symbols constitutes a word can be discussed. Dale (1965, p.898) described the range of knowing a word. On a continuum, one may 1) never have seen the word, 2) know there is such a word but not know its meaning, 3) have "a vague contextual placing of the word.", 4) recognize it again and remember it, or 5)require very fine distinctions, e.g. between "sympathy and empathy." Dale believed the nature of the test--that is, where on the continuum a word is to be defined--influenced the size of vocabulary, and thus, was a source of estimated size differences.

Dale (1965, p.897) was also concerned by the lack of assessing depth of meaning. He stated, "One of the major deficiencies of word lists which are supposedly concerned with meanings, not spelling, is that they have not usually studied frequency of meanings." Then he pointed out, however, that the difficulty was met by A Semantic Count of English Words, published in 1938 by Thorndike and Lorge. This list gives the frequency with which different meanings

of words in print occur. Lorge and Chall (1963) were similarly concerned and ask whether one has to know all meanings of a word or only the commonest meaning to "know" the word.

Lorge and Chall (1963) also raised questions about testing knowledge of a word. To indicate that a word is known, one may define it, use it in a sentence, check it on a list, or pick it from four or five choices. Multiple choice tests offer the convenience of short administration time, ease of correction, and relative inexpense. But there are some difficulties in this method, mainly in the possible combinations of knowing the key and the choices. A subject may know the key, but not in terms of the choices, or may not know what the choices mean. The subject also may not know the key, but know all the choices. One is given credit for knowing only the key word. Lorge and Chall questioned whether one should not also be given credit for knowing the correct alternative.

Another problem in testing word knowledge is the need to decode words on a reading test. A subject can fail a word because he is not able to decode the key or the alternatives though the meanings are known. Oral administration alleviates this problem. Lorge and Chall (1963) suggested giving a test orally to those people who can not recognize a word in print, if the word <u>meaning</u> is being measured. Since many children in primary grades have difficulty decoding words, oral administration would allow measurement

of word understanding. However, the task becomes measurement of understanding vocabulary, not reading vocabulary.

Separation of the reading process from word understanding was further expressed by Jackson (1973, p. 9). He was concerned with what vocabulary development is about. He stated,"...when we are teaching vocabulary we are attempting to foster in the youngster an increasing knowledge of his world and to develop the representation of this knowledge through verbal symbols or words." Much of this development occurs orally and Jackson feels that once a meaning is represented by an oral symbol, recognizing the visual symbol is irrelevant to vocabulary development. The child may have a problem in word naming, but not in vocabulary development.

Designation of vocabulary. Another area that causes confusion is the failure to state which vocabulary is being measured. We have listening (aural understanding), speaking, writing, and reading vocabularies. These are not the same size, and their relative size toward each other changes with age. At first one's reading and writing vocabularies are very small compared to his speaking and aural understanding vocabualry. With age, one's reading and writing vocabularies overtake speaking and aural understanding vocabularies. Lorge and Chall (1963, p.148) stressed the need for stating which vocabulary is being measured for young children:

> Particularly for first-grade children whose reading and writing vocabularies are first being developed, it is essential that estimates be understood in terms of the method employed--whether based on words spoken, understood when heard, or understood when read in print.

When Rinsland (1945) found first-graders knew about 5,000 words, it must be remembered this was writing vocabulary size. Ames (1964) found first graders knew 12,456 words, which was the mean understanding vocabulary size. Lorge and Chall (1963, p.148) cite Neher's study: "Estimates based on speech or writing of adults are usually lower than those obtained from sampling methods, since a person's recognitional vocabulary is considerably greater than his expressional vocabulary..."

Errors in sampling. One cause of variations in vocabulary size estimates is errors in sampling frequency counts. In general, these errors tend to limit the estimated size of one's vocabulary. Dale (1965, p.897) made the following criticism of frequency-sampling: "Investigators miss simple words which may not appear in print but are well-known. Common words are often omitted. Rinsland, for example, has "platypus" but not "pimple", "mucus" but not "snot"..."

Errors in sampling dictionaries can either limit or inflate vocabulary size. Problems in using this method include dictionary size, editors' selection of words to include in a dictionary, and the space allocated to the words (Lorge and Chall, 1963). There is also a problem of recen-

cy. The dictionary which is being sampled limits the size of one's vocabulary to the size of that dictionary. A small dictionary may not have some of the words a person knows whereas a large unabridged dictionary has a greater chance of having included these words.

Colvin (1951) cites several studies using abridged and unabridged dictionaries. The early studies utilized abridged dictionaries. A sample was taken with each word representing 10 or 20 words. Terman found twelve-year olds knew 7,200 words. The same list of words was used, but a different type of test given, in which Holley found twelve-year olds knew 8,478 words. When Kirkpatrick tested seventh-graders on a list of 100 words from Webster's Academic Dictionary, they knew 10,666 words. Later studies used unabridged dictionaries. Gerlach found college sophmores knew 85,300 words, while Seashore, using a different dictionary, found college sophomores knew about 70,000 words. M.K. Smith found seventh-graders knew 55,000 words. These figures support Hartmann's (1941) findings that vocabulary estimates depend on the size of dictionary sampled.

Another problem Lorge and Chall (1963) point out is the editors' selection of words to include in a dictionary. An abridged dictionary has a biased sample of the more common words from the total English vocabulary. In unabridged dictionaries, editors may have cut old or uncommon words. Thus, even a sample from a large unabridged dictionary is not a sample of the universe of the English language.

While old words may have been cut, what about new words being added? Little mention was made in the literature about the recency of the dictionary being sampled. Studies that depend on the Seashore-Eckerson test (or any test dependent on a sample from an earlier edition dictionary) ignore the increase in vocabulary from World War II through present space exploration.

Lorge and Chall (1963, p.152) cite Thorndike on an error in dictionary-sampling which has to do with the space-sampling of words. "...selecting words falling in a particular place in the dictionary, for example the first word on a page would overweight the sample with the most important words, since most dictionaries give more space to important words."

Another cause of variation in vocabulary size, when using the dictionary-sampling method, is the representativeness of the sample. The Seashore-Eckerson Test was critiqued by Lorge and Chall (1963, p.150-153). The major criticism was that Seashore and Eckerson had a biased sample of easier words, since they did not adjust for multiple entries of words. Also, their basic words included names of people, geographical names, word parts, proper nouns, and duplicate spellings, along with the several entries of the same word form. Counting multiple entries was criticized because knowledge of the word was determined by recognition of only one common meaning of the word. ("Cow" may have been listed as eight words, but if a student knew

it was an animal, and no other definitions, he would still have been credited with knowing eight words.) Studies that were based on the Seashore-Eckerson Test, e.g. Colvin (1951), Bryan (1953), and Shibles(1959), also had the results held in question. It was concluded that all these studies had inflated vocabulary size estimates.

Tests

Most investigators in the literature constructed their own tests for measuring vocabulary size. Other than vocabulary sections in battery tests, there are few tests designed specifically to measure vocabulary size of young children. Tests in Print II (Buros, 1974) has a list of 20 tests that measure a range of types of vocabulary. These included tests for business and professionals, tests appropriate for ages 18 and up, or appropriate for grades nine and up.

Three tests appropriate for grades three and up were reviewed in Buros' Mental Measurements Yearbooks. The first was The Survey Test of Vocabulary (Buros, 1949, test 167), based on Thorndike's Teacher's Word Book of Twenty Thousand Words. It was more than 35 years old. No information was furnished concerning test construction, or when the norms were obtained. The second test was The Vocabulary Test: National Achievement Tests (Buros, 1949, test 168) which was more than 30 years old. It had no information on word selection other than "...the words used have

been carefully selected and graded." The third test was the Wide Range Vocabulary Test (Buros, 1949, test 169) which also was more than 30 years old. There were no data on reliability, intercorrelation between forms, or correlation with other well known measures.

There are foreign language tests (English as a second language) available which measure English vocabulary. These tests were designed for older students and tended to check idiomatic expressions, however.

The test developed by Dupuy (1974) is quite a contrast to the ones previously mentioned, in terms of word definition, test construction, and standardization. The Basic Word Vocabulary Test is discussed in chapter Three under Instrument.

Summary

Controversy over estimated vocabulary size has continued since investigation began. Lorge and Chall (1963) pointed out methodological issues that contributed to varying vocabulary size estimates. Although difficulties are present, the dictionary sampling method appears to represent the most words that a person could know. The problems of defining a word, defining know, and sampling have not been completely settled or eliminated. Therefore, there is need to define each aspect of measurement in any study undertaken, so the reader can be aware and interpret results in light of these issues.

CHAPTER III

PROCEDURE

This study was an attempt to extend the Basic Word Vocabulary Test (BWVT) downward by administering it orally. Some investigations included measuring the aural understanding basic word vocabulary size of primary grade children; finding correlations between oral scores and IQ, between oral scores and reading scores; and determining if significant differences exist between the oral BWVT and the reading BWVT scores. This chapter is a description of the population of the study, instrument, the pilot study, research design, data collection, and data analysis.

Population

All the subjects in this study live in the township of Green Brook, New Jersey, and attend its one elementary school. Ninety-six subjects participated in the experiment with an equal number of subjects (24) being randomly selected from grades one through four. It was felt this number of subjects would be manageable in terms of administering and scoring the tests. Half the subjects were girls, half were boys.

According to the 1970 U. S. census, the township in

which the subjects live has 3,838 inhabitants. Of the people who were 25 or older, 44% were high school graduates. The median income was $11,148, and mean income was $12,128. The labor force was comprised of 1,660 people 16 years and older; 64% were in sales, clerical, craftsmen, operatives,and transport positions. All subjects were white, and either Catholic or Protestant.

All subjects' IQ's were measured when they were in the second grade. Each year in November, the Kuhlman-Anderson Test had been given by the second-grade teachers. (Thus, there were no IQ data for children in first grade.) In his review of the Kuhlman-Anderson Test, Michael (Buros, 1965) said it "compares favorably with any other competitive instrument measuring general intelligence."

The second-grade subjects' IQ's ranged from 99 to 140 with a mean of 117, the third-grade subjects' IQ's ranged from 104 to 133 with a mean of 118, and the fourth-grade subjects' IQ's ranged from 98 to 140 with a mean of 119. Fourth graders took the Cognitive Abilities Test in February 1976. Their scores ranged from 81 to 140, with a mean of 108. This test was not reviewed in Buros' Mental Measurements Yearbooks.

Reading abilities were measured by the Gates-MacGinitie Reading Test, Primary Forms A, B, and C. In his review of this test, Van Roekel (Buros,1970) found the standardization "to have been rather carefully done" but when "used alone, however, these tests probably function

best as survey tests." Since the tests were given in May, '75, reading abilities would be higher than indicated due to passage of time and additional instruction. The second graders' vocabulary scores ranged from 1.4 to 3.5, while the third graders' vocabulary scores ranged from 2.2 to 5.2. The fourth graders' vocabulary scores ranged from 3.0 to 6.8. No standardized test information on reading ability was available for first graders.

Instrument

Dupuy's Basic Word Vocabulary Test (BWVT) measures the absolute and the relative size of a literate person's reading vocabulary. Its range extends from third grade through Ph.D. The BWVT has high reliability and validity as a reading test. The overall internal consistency reliability is close to .96. Its standard error of measurement is 3.0 raw score points (Dupuy, 1974, p.30).

For concurrent validity, the BWVT scores were correlated with scores on the School and College Ability Tests (SCAT), scores from five different tests of the Sequential Tests of Educational Progress (STEP), and with test scores from verbal sections of five standardized tests. These five standardized tests included the Metropolitan Reading Readiness Test, the Lorge-Thorndike Intelligence Test, the California Test of Mental Maturity, the Differential Aptitude Test, and the School and College Ability Tests. A median correlation of .76 was found. The correlations

ranged from .60 to .84 at grades 4 to 12, but were .28, .45, .46 at first, second, and third grades, respectively. Fourth grade correlation was .80. Dupuy (1974, p.30) states, "The low correlations for grades 1, 2, 3 are consistent with the findings that the BWVT is too difficult at these levels when given as a reading test." This is why this study is being done with oral administration.

Basic word knowledge is highly related to age and educational levels. Correlations of the BWVT and educational level were .80 for grades 3-12 and .60 for grades 3-7. BWVT and age level correlations were .77 for 8-17 year olds and .55 for 8-12 year olds at better than the .01 level (Dupuy, 1974, p.31).

The content validity can be evaluated by reviewing the procedures Dupuy (1974, p.4-6) used to select the Basic Word Vocabulary. Criteria for drawing a sample of main entry words from Webster's Third New International Dictionary were prepared. The dictionary has 450,000 entries. Main entry words had to meet these criteria: 1) the word had to be a main entry, 2) all homographs were counted as one word, 3) abbreviations were counted, but not affixes, and 4) letters were not counted.

An estimate of the total number of main entry words was made following explicitly set procedures. Basically, every 10th page had one column sampled. Computations yielded a rounded population estimate of 240,000 words. A 1-percent sample (2,400 words) was taken by selecting the

third word (same criteria applied as before) in one column on every page. The order of the columns was rotated. These 2,400 words were put into one of four categories: 1) compounds or hyphenated, 2) proper names, 3) abbreviations, 4) others. Then, only the 1,360 "others" were checked against main entries in Random House, World Book, and Funk and Wagnall's. 381 words were main entries in all the four dictionaries. These words were reduced by eliminating technical, foreign, slang, and archaic words, as classified by Random House.

Another set of criteria was developed to classify the remaining 307 words as either basic or derived. A derived or variant entry word was 1) another form or different tense of the same word, 2) a different spelling, 3) another word which gave a fuller definition, 4) a two or more word entry that used one of the words in the definition, and 5) a derived word whose meaning could be known if the base and affix meaning were known. A sample of 123 words resulted which have been carefully and precisely defined. Logically, the test has excellent content validity.

The test (of the 123 basic words) is a multiple choice test that has five choices for each item--one correct response and four distractors. Similarly precise criteria for constructing the correct responses and distractors were developed. For example, correct responses had to be the most common meaning (according to World Book Dictionary. They had to be easier (more frequently used) than the stem

word, and usually were a one word synonym. Alliteration similarity to the stem was avoided. Correct responses were balanced in length to distractors, and in parallel form in tense or part of speech. Distractors had to be less difficult than the stem, in parallel form to the stem and other choices, and alliteration and spelling similarities were avoided.

When the test construction was completed, two pretests were given. About 50 of the pretest subjects evaluated the items. Then, standardization was done on 3,100 students in grades 1-12 in the Fairfax County, Virginia public schools. The parents represented a diverse background of social status, occupation, and geographic origin. Item analyses were made to find difficulty level, internal consistency, distractor effectiveness, and sex differences for each item (Dupuy,1974,p.11). By using the nationally standardized test scores of verbal achievement, grade and age norms were developed and adjusted to reflect national norms.

Upon administration of the test, a vocabulary size estimate can be easily calculated by multiplying a subject's raw score by 100. There are 123 basic words in a one percent sample of the main entries in Webster's Third New International Dictionary. So, multiplying the number correct out of 123 by 100 yields an estimated basic word vocabulary size for an individual. In arriving at a figure for vocabulary size, it must be remembered that the unit of measurement is a basic word, and that the vocabulary being

measured is an aural understanding vocabulary.

In this study, the first section of the BWVT was used for oral administration. Obviously, since the test's range is through Ph.D. level, primary grade children would not have to take the entire test.

Pilot Study

The pilot study for this experiment was done in February, 1976. Five children per grade level (kindergarten, first, second, third, and fourth grades) were randomly selected to participate in the study. At each grade level, five names were drawn from a box. Each group of five was given the BWVT orally. The study aided in foreseeing difficulties that would arise in running the actual test. For example, kindergarteners could not handle marking an item with five alternative choices in a group situation. Perhaps a different format for the test or individual administration would work at this grade level.

Design

It was felt that 24 students per grade level would be a manageable number in terms of administration and scoring. Therefore, 96 students in first through fourth grades were randomly selected for the Basic Word Vocabulary Test (BWVT) administration. Those students who participated in the pilot study were not included in the sample who took the test. At each grade level, all students' names, except

those who had participated in the pilot study, were put in a box. A stratified random sample was selected by taking the first 12 girls' names and the first 12 boys' names at each grade level. Investigation of sex differences could be easily undertaken, if desired at a later date.

Students in third and fourth grade were given the BWVT first as a reading test and one week later, as an oral test. The dual administration provided data to compare reading and oral vocabularies. Students in first and second grades had oral administration only.

Data on age, IQ, and reading scores of subjects were collected from school files.

Data Collection

Subjects were tested during February and March of 1976. The BWVT was administered in the afternoon, with each group using approximately 35 minutes to take the test. Testing was done in groups of 4, 12, or 24. The first-graders and second-graders were tested in groups of 4 at a time. The small number facilitated administration; the examiner could make sure each student was on the current question. The third-graders were tested 12 at a time; the fourth graders were tested 24 at a time.

The test was administered only by the experimenter for all grades. Children were moved to the experimenter's classroom for testing. The subjects were told the examiner was working on a project (like they did in school) to find

out how many words children of different ages knew. They were assured that this test would not affect their report card grades and that the results would only be seen by the examiner. Test booklets were distributed and directions read to the subjects--they were instructed to mark an X next to the correct choice on each item. The examiner read each item (stem and five choices) aloud.

The BWVT was not given in its entirety. Since the items were ranked according to difficulty, children through fourth grade would not know words at the upper end. Enough items had to be administered to ensure the highest-scoring subject would get 10 errors. The highest scores from the pilot study helped estimate a cut-off point. If a subject had not made 10 errors, he was tested further.

Scoring was done by Dupuy's 10-E method. In this method, a subject's raw score is determined by counting the number correct, and counting is stopped at the 10th error. Omitted items are not included in the count.

Data Analysis

For each grade level, the means and standard deviations from the oral administration of the BWVT were calculated. Then the means were multiplied by 100 to get the estimated understanding vocabulary size of basic words.

To verify significant differences between first and second grades, second and third grades, and third and fourth grades, an analysis of variance and Scheffe's tests

of contrasts were run.

The means and standard deviations of the reading administration of the BWVT were calculated for the third grade and for the fourth grade. To get an estimate of the third grade and the fourth grade reading vocabulary sizes, the means were multiplied by 100.

To determine if significant differences existed between oral and reading scores at the third grade and at the fourth grade levels, two *t* tests were made.

A correlation matrix was run to find correlations between the oral BWVT and 1) the Gates-MacGinitie Primary Reading (vocabulary) Test, 2) the Gates-MacGinitie Primary Reading (comprehension) Test, and 3) the Kuhlman-Anderson IQ Test. The statistics shown used Pearson product-moment correlations.

CHAPTER IV

FINDINGS AND DISCUSSION

This chapter presents an analysis of the data obtained from raw scores on the Basic Word Vocabulary Test (BWVT), given orally to 96 subjects in first through fourth grades. Data also analyzed are the scores from the BWVT given as a reading test to the third and fourth graders.

Presentation of Data

The data from the oral administration of the BWVT, the reading administration of the BWVT, and data collected from students' school files are presented in Table 2. The mean raw scores on the oral BWVT were 17.8 (S.D.=5.4), 23.3 (S.D.=4.6), 37.8 (S.D.=11.2), and 40.0 (S.D.=13.2) for the first, second, third and fourth grades respectively. When the means were multiplied by 100 to produce the aural understanding vocabulary size estimate, it was found that first graders knew 1,780 basic words, second graders knew 2,330 basic words, third graders knew 3,780 basic words, and fourth graders knew 4,000 basic words.

The mean reading scores on the BWVT were 27.6, and 31.3 (S.D.=12.6) for the third and fourth grades respectively.

TABLE 2

MEANS FOR ORAL BWVT, READING BWVT, VOCABULARY SIZES, AGES, IQ, AND GATES-MACGINITIE READING TEST BY GRADE LEVEL
(N = 96)

GRADE[a]	BWVT ORAL			READING BWVT			CA	MA	IQ[b]	GATES GRADE LEVEL[c]	
	RAW SCORE[d]	SD	VOC. SIZE	RAW SCORE	SD	VOC SIZE				VOC.	COMP.
1	17.8	5.4	1780				82.0				
2	23.3	4.6	2330				94.6	110.5	117.2	2.7	2.7
3	37.8	11.2	3780	27.6	11.4	2760	105.7	124.8	118.2	4.3	4.1
4	40.0	13.2	4000	31.3	12.6	3130	117.8	138.5[e]	119.1[e]	5.0	4.7

[a] N = 24 for each grade.

[b] Data are one year old for third grade and two years old for fourth grade.

[c] Data are old--add .6.

[d] Data are significantly different by analysis of variance and Scheffe's tests on contrasts.

[e] Mean is based on only 20 scores, since 4 of 24 subjects had no IQ data.

The raw scores were multiplied by 100 to find the reading vocabulary size estimates. Third graders knew 2,760 basic words, while fourth graders knew 3,130 basic words.

Data from school files include students' ages, IQ, and Gates-MacGinitie Reading Test results. In this school system, Kuhlman-Anderson IQ tests were administered only during second grade. Therefore, no data for first grade were available, and the data for third and fourth graders were old. The mean IQ's for second, third, and fourth grades were 117.2, 118.2, and 119.1 respectively. The mean IQ of the fourth grade is based on 20 scores only, since four subjects had no IQ data. Chronological age (CA) is calculated for age in months at testing time (February). (The grade level is actually an extra .5 since testing occured in February.) Mental age (MA), also expressed in months, was determined by using the equation MA = CA X IQ/100. Reading test scores were from the previous May, so students' actual reading levels would be higher (add .6 for six months) due to passage of time and additional instruction. Mean Gates' vocabulary grade levels for second, third, and fourth graders were 2.7, 4.3, and 5.0 respectively. Mean Gates' comprehension grade levels for second third, and fourth graders were 2.7, 4.1, and 4.7 respectively.

Significant Differences

It is expected that each grade level mean score on the

BWVT should be higher than the mean score of the preceding grade. To verify that the differences were significant, a one way analysis of variance was applied to the mean raw score results on the oral BWVT. The analysis revealed a main effect of grade ($F = 32.06$, $df = 3/92$, $p < .01$), shown in Table 3. Further analysis of the main effect means, using Scheffe's tests on contrasts, indicated that all pairwise comparisons were significant ($p < .01$), shown in Table 4.

To determine if differences between the oral BWVT scores and the reading BWVT scores were significant, *t* tests were performed between the third grade oral and reading means and between the fourth grade oral and reading means. Results are shown in Table 5. The *t* test performed on the third grade oral raw score means and the reading raw score means was 6.9, significant at the .001 level. The *t* test performed on the fourth grade oral raw score means and the reading raw score means was 6.3, also significant at the .001 level. The problem of results being confounded by the practice effect will be taken up in the following Discussion section.

Correlations

To find the correlations between the oral BWVT scores and IQ, and between the oral BWVT and the Gates-MacGinitie Reading Test sections, a correlation matrix was calculated.

TABLE 3

ANALYSIS OF VARIANCE AMONG GRADE LEVELS

(N = 96)

SOURCE OF VARIATION	SUM OF SQUARED DEVIATIONS	DF	MEAN SQUARE	F RATIO
AMONG SAMPLES	8457.06	3	2819.02	32.06**
WITHIN SAMPLES	8088.81	92	87.92	
TOTAL	16545.88	95		

** F (3/92) = 32.06; p > 0.01

TABLE 4

ANALYSIS OF MEAN DIFFERENCES AMONG GRADE LEVELS USING SCHEFFE'S TESTS ON CONTRASTS (N = 72)

GRADES	1	2	3	4
1	-	5.542**	19.958**	22.167**
2			14.417**	16.625**
3				2.208**
4				-

**p < 0.01

TABLE 5

t TEST BETWEEN ORAL ADMINISTRATION AND READING ADMINISTRATION OF THE BWVT FOR THIRD AND FOURTH GRADES (*N* = 48)

GRADE[a]	MEAN ORAL RAW SCORE	MEAN READING RAW SCORE	*t* VALUE	SIGNIFICANCE LEVEL
3	37.8	27.6	6.9	0.001
4	40.0	31.3	6.3	0.001

[a] *N* = 24 for each grade.

There were 68 subjects who had complete information for oral BWVT, Gates reading vocabulary, Gates comprehension, and the Kuhlman-Anderson I.Q. Four fourth graders did not have IQ data. The results are shown in Table 6. The oral BWVT has a .74 r with the Gates (reading) vocabulary, a .59 r with the Gates (reading) comprehension, and a .29 r with the Kuhlman-Anderson IQ test.

TABLE 6

CORRELATION MATRIX OF ORAL BWVT RAW SCORES, GATES VOCABULARY GRADE SCORES, GATES COMPREHENSION GRADE SCORES, AND KUHLMAN-ANDERSON IQ SCORES (N = 68)[a]

	BWVT	GATES VOC.[b]	GATES COMP.[b]	K-A IQ[c]
BWVT	-	+.74**	+.59**	+.29*
GATES VOC.			+.89**	+.40**
GATES COMP.				+.44**
K-A IQ				-

[a]Data complete for second, third, and fourth grades only.

[b]Data were six months old for all grades.

[c]Data were one year old for third graders and two years old for fourth graders.

*$p < 0.05$

**$p < 0.01$

Discussion

The first question to be examined was the size of children's understanding vocabularies in grades one through

four. The results (1,780 for first grade, 2,330 for second grade, 3,780 for third grade, and 4,000 for fourth grade) were much lower than the results of previous studies. This is rather startling in light of the IQ and reading level of this experiment's population. One could expect students from a more normally distributed IQ population to score even lower. The low number of words is not due to students doing poorly on the test. The instrument measures a maximum basic word vocabulary of 12,300 words, which is Ph.D. level size of vocabulary. In the Ames (1964) study, 12,456 words was the estimated vocabulary size of first graders. The issues raised by Lorge and Chall (1963) indicate reasons for the low vocabulary size estimate.

The first methodological issue, raised by Lorge and Chall (1963), in the dictionary sampling method, was dictionary size. The low estimates were not due to using a small, abridged dictionary. The sample of main entry words was drawn from Webster's Third New International Dictionary which contains more than 450,000 entries.

The second issue discussed by Lorge and Chall(1963) was the definition of a word. The criteria for the unit of measurement--the basic word--seem to be responsible for the low vocabulary size result. Dupuy (1974, p.5,6,7) estimated the word population of Webster's Third New International Dictionary to be 240,000 main entries. He reduced the one-percent sample of 2,400 words to just 123 basic words by eliminating words classified as extra homographs,

compounds, abbreviations, hyphenated, proper nouns, technical, foreign, slang, archaic, derived, variant, or redundant. The basic words also had to be present in three other large specified dictionaries.

In comparing the procedures Ames (1964) used, there are fewer and less strict criteria for defining which words should be included for testing. Ames sampled the Funk and Wagnalls New Standard Dictionary (1956). Using a table of random numbers, he chose 500 words that were main entries. Four affixes and three abbreviations were eliminated, leaving 493 words. No mention was made about compounds, hyphenated forms, or proper names. No check was made to see if the 493 words were main entries in other dictionaries. Since the test was to be administered to first graders, Ames had five graduate students eliminate difficult words. There remained 181 words, which Ames tested on ten sixth graders. Ames (1964, p.66) "asked for responses until he was sure that at least one member of the group knew an appropriate meaning for the word." Neither set of students seemed to have specific criteria for eliminating words. Probably some foreign, slang, technical, and archaic words were eliminated, but Ames does not specifically state this. Nothing was said about derived, variant, or redundant forms. Dupuy started with a one percent sample of 2,400 words and reduced it to 123. Ames started with a 500 word sample and reduced it to 121. Dupuy's 123 words range from first grade through Ph.D. level, while Ames' 121 words were

all meant for first graders.

The second question to be examined was the existence of significant differences among the grade levels. The results of the analysis of variance shown in Table 3 and Scheffe's tests on contrasts shown in Table 4 verify that there were significant differences. It is interesting to note that there is not such a great difference between grades one and two, or between grades three and four, but there is a large difference between grades two and three. The difference could be due to a growth spurt with a subsequent plateau, or it could be that particular items were difficult for the fourth graders. A broader range study and/or an item analysis would give additional information on this question.

The third question in the study dealt with the existence of significant differences between oral and reading scores for the third and for the fourth grades. While the results of the *t* Tests provide evidence for reliable differences, the results may have been confounded by the practice effect. Both grades were administered the reading version first, then, one week later, the oral version of the BWVT. For each grade level, the oral scores could be higher due to practice alone, rather than an actual larger sized oral understanding vocabulary. This effect could have been avoided by administering the BWVT as a reading test first to half of a grade, and as an oral test first to the other half.

The last question of the study was concerned with the existence of correlations between the oral BWVT, Gates Reading Test results, and IQ. The results of the oral administration of the BWVT correlate moderately (r = .74) with the Gates reading vocabulary for grades two, three, and four, but there is a lower correlation (r = .59) between the BWVT and the Gates comprehension section. One might expect the correlation between BWVT and comprehension to be lower than the correlation between vocabulary and BWVT. There is a discrepancy in time in these calculations--the Gates' test had been given the preceding May, so .6 could be added to the reading levels before performing the correlations. Perhaps using a different type of standardized reading vocabulary test (or vocabulary section) would also yield higher correlations. The Stanford Diagnostic Reading Test (upper elementary level) measures vocabulary (in stanines) by having children choose one of three orally pronounced choices that best completes an orally pronounced sentence. This procedure is a closer approximation of the oral administration of the BWVT.

There is a low correlation with IQ (r = .29, $p < .05$). This is an unusual result because, generally, vocabulary tests and IQ do correlate highly. Again, the time of IQ test administration must be taken into account. While the IQ data are three months old for second graders, the data are one year and two years old for third and fourth graders respectively. The correlation may be higher with more

current IQ data.

Since the BWVT was administered as a reading test to the third and fourth graders, the results can be directly compared to Dupuy's study. The raw score means for Dupuy's third and fourth graders were 15.3, and 27.0, respectively, while the reading raw score means from this study were 27.6 (third grade) and 31.3 (fourth grade). Each grade in the two studies was tested midway through the academic year, so differences between studies were not due to extra instructional time. The chronological-age grade placement of the students in this study was 3.6 for third graders and 4.6 for fourth graders, indicating they were within the normal range. However, the mental-age grade placement was 5.2 for third graders, and 6.2 for fourth graders, indicating this sample had above average ability.

This study showed the feasibility of administering the BWVT orally. Even though the sample size was small (N= 96), and the data from student files was two years old for some subjects on the correlations, significant information about vocabulary size was found.

CHAPTER V

SUMMARY AND CONCLUSIONS

This chapter summarizes the study, draws conclusions from the research results, and offers suggestions for additional research on vocabulary size.

Summary

This study was an attempt to extend the Basic Word Vocabulary Test (BWVT) downward through oral administration. The first question to be investigated in the study was, "How large are the aural understanding basic word vocabularies of children in grades one through four?" In order to answer this question, 96 subjects from grades one through four were randomly selected to participate in the study. The subjects went to the one township school in a Central New Jersey suburban community. The BWVT was administered as a reading test first to third and fourth graders, and one week later, the BWVT was given to them orally. The first and second graders had oral administration only, and they were tested in groups of four at a time. The raw scores were converted into means by grade level and multiplied by 100 to estimate aural vocabulary sizes. First graders knew 1,780 basic words, second graders knew 2,330

basic words, third graders knew 3,780 basic words, and fourth graders knew 4,000 basic words.

The second question to be investigated was, "Do significant differences exist among grade levels?" A one way analysis of variance showed main effect for grade (F = 32.06, df = 3/92, p <.01). A further analysis using Scheffe's tests on contrasts showed all pairwise comparisons were significant (p < .01).

The third question to be studied was, "Do significant differences exist between the oral BWVT scores and the reading BWVT scores of third and fourth grade children?" A t test was performed between the third grade oral BWVT mean (37.8) and the reading mean (27.6). The resulting t value of 6.9 was significant at the .001 level. A t test was also performed between the fourth grade oral BWVT mean (40.0) and the the reading mean (31.3). The resulting t value of 6.3 was also significant at the .001 level. Results may have been confounded due to the practice effect.

The last question to be investigated was, "What correlations exist between the oral scores and IQ, and between oral scores and the Gates-MacGinitie Reading Test?" A moderate correlation of .74 between the oral BWVT and the Gates vocabulary test was found, while a lower correlation of .59 between the BWVT and the Gates comprehension test was found. The Kuhlman-Anderson IQ Test had a very low correlation of .29. There is a discrepancy of time in the

data taken from students' files. The reading test data is six months old, and the IQ data is one year old for third graders, and two years old for fourth graders. The correlations may actually be higher than the results indicate.

Conclusions

Many problems exist in vocabulary size research. The BWVT is an instrument that overcomes one of the biggest difficulties--the definition of the unit of measurement--the basic word. However, because basic words have to meet several criteria, there are less of them in proportion to the total word population estimate. This makes for a lower vocabulary size estimate than the results of previous studies.

In this study, the major conclusion is that the BWVT can be extended downward through oral administration. However, in its present format, it is best administered to groups of four or five at the first and second grade levels. Vocabulary size at each of the grade levels was lower than previous estimates, such as Ames (1964). To reiterate, the low vocabulary size is probably due to the strict criteria in defining the unit of measurement--the basic word.

The results provide evidence for the importance of vocabulary teaching. The children came to school knowing a certain amount of words, and increased their vocabularies significantly from year to year while in school. There was

no special emphasis on vocabulary study, so a curriculum that would emphasize vocabulary study might produce students who would score within the next higher standard deviation.

From Dupuy's study, it can be seen that the total number of basic words to learn is relatively small, compared to the total number of dictionary entries. If this is true, students would know more "words" if they knew some basic words, and a few rules for adding affixes, making compounds, and variant forms. Additional study in this area could lead to some curriculum change.

Suggestions for Further Research

The first suggestion for further study is to replicate this experiment, using a larger number of subjects. The subjects should be of a more normally distributed IQ population. One should expect the results to be lower vocabulary size estimates, in that case.

Another suggestion is that, since the BWVT yields an absolute vocabulary size, it could be used to measure wide range vocabulary growth. A longitudinal study on individuals over a two-year, three-year or longer period would give results indicating how much growth they made. If the study were done over different grade levels, trends in rate of vocabulary growth could be monitored.

The reading vocabulary size is larger than the understanding vocabulary size in adults. At what point the

reading vocabulary becomes larger than the oral understanding vocabulary could be studied. The BWVT could be administered orally and as a reading test to several grade levels above fourth to find the crossover point. Care must be taken to avoid the practice effect: each grade could be divided in half, with one half taking the oral version first and the other group taking the reading version first.

BIBLIOGRAPHY

American Psychological Association. Publication Manual (2nd ed.). Washington, D.C.: American Psychological Association, 1967.

Ames, W. S. The understanding vocabulary of first-grade pupils. Elementary English, 1964, 41, 64-68.

Bryan, F. E. How large are children's vocabularies? Elementary School Journal, 1953, 54, 210-216.

Buckingham, B. R., & Dolch, E. W. A combined word list. New York: Ginn, 1936.

Buros, O. K. (Ed.). The seventh mental measurements yearbook. Highland Park, N.J.: Gryphon Press, 1972.

Buros, O. K. (Ed.). The sixth mental measurements yearbook. Highland Park, N.J.: Gryphon Press, 1965.

Buros, O. K. (Ed.). The third mental measurements yearbook. New Brunswick, N.J.: Rutgers University Press, 1949.

Buros, O. K. (Ed.). Tests in print II. Highland Park, N.J.: Gryphon Press, 1974.

Colvin, C. A re-examination of the vocabulary question. Elementary English, 1951, 28, 350-356.

Dale, E. The problem of vocabulary in reading. Educational Research Bulletin, 1956, 25, 113-123.

Dale, E. Vocabulary measurement: Techniques and major findings. Elementary English, 1965, 42, 895-901+.

Dale, E. & Razik, T. Bibliography of vocabulary studies (4th ed.). Columbus, Ohio: Ohio State University, 1963.

Dolch, E. W. How much word knowledge do children bring to grade 1? Elementary English Review, 1936, 13,

177-183.

Dolch, E. W. Methods in reading. Chicago: Garrard Press, 1955.

Dupuy, H. J. Basic word vocabulary test. Highland Park, N.J.: Drier Educational Systems, 1975.

Dupuy, H. J. The rationale, development, and standardization of a basic word vocabulary test (DHEW Publication No. (HRA) 74-1334). Washington, D.C.: U.S. Government Printing Office, 1974.

Gerlach, F. M. Vocabulary studies--false definition test. Colorado College Studies in Education and Psychology, 1917, 1.

Hartmann, G. W. Crtique of the common method of estimating vocabulary size. Journal of Educational Psychology, 1941, 32, 351-358.

Holley, C. E. Holley sentence vocabulary scale--grades 3-12. Bloominton, Ill.: Public School Publishing Company, 1919.

Horn, M. D. (Chairman). A study of the vocabulary of chilren before entering the first grade. Washington, D.C.: International Kindergarten Union, 1928.

Jackson, R. K. Does the world need another article on vocabulary development? Reading Improvement, 1973, 10 (1), 9-11.

Jacobs, H. D. A replicative investigation of the Buckingham-Dolch free-association word study. Eugene, Oregon: Oregon University, 1967. (ERIC Document Reproduction Service No. ED 017 333)

Kirkpatrick, E. A. Vocabulary test. Popular Science Monthly, 1907, 70, 157-164.

Kolson, C. J. Oral arithmetic Vocabulary of kindergarten children. The Arithmetic Teacher, 1963, 10, 81-83.

Larrick, N. How may words does a child know? Education Digest, 1954, 19, 42-44.

Lorge, I. & Chall, J. Estimating the size of vocabularies of children and adults: An analysis of methodological issues. Journal of Experimental Education, 1963, 32 (2), 147-157.

Rinsland, H. D. A basic vocabulary of elementary school children. New York: Macmillan, 1945.

Seashore, R. H. The measurement and analysis of extent of vocabulary. Psychological Bulletin, 1933, 30, 709-710.

Seashore, R. H., & Dolch, E. W. Implications of the Seashore vocabulary report: with reply. Elementary English, 1949, 26, 407-413.

Seashore, R. H., & Eckerson, L. D. Measurement of individual differences in general English vocabulary. Journal of Educational Psychology, 1940, 31, 14-38.

Seegers, J. C. Vocabulary problems in the elementary school: A digest of current research. Elementary English, 1939, 3, 155-166.

Seegers, J. C., & Seashore, R. H. How large are children's vocabularies? Elementary English, 1949, 26, 181-194.

Shibles, B. H. How many words does a first grade child know? Elementary English, 1959, 31, 42-47.

Smith, M.E. An investigation of the development of the sentence and the extent of vocabulary in young children. University of Iowa Studies in Child Welfare, 1926, 3 (5), 1-92.

Smith, M. K. Measurement of the size of general English vocabulary through the elementary grades and high school. Genetic Psychology Monographs, 1941, 24, 311-345.

Terman, L. M. The measurement of intelligence. Boston: Houghton Mifflin, 1916.

Wesman, A. G., & Seashore, H. G. Frequency versus complexity of words in verbal measurement. Journal of Educational Psychology, 1949, 40, 395-404.

APPENDIX A

SAMPLE BASIC WORD VOCABULARY TEST

(removed due to copyright restrictions)